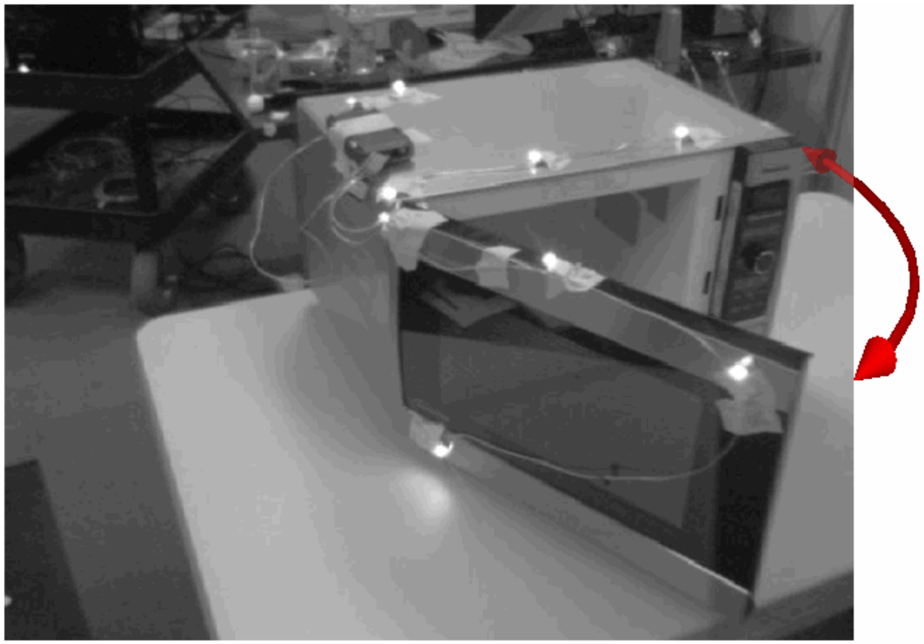
Towards Understanding Articulated Objects



Jürgen Sturm[1]
Cyrill Stachniss[1]
Vijay Pradeep[2]
Christian Plagemann[3]
Kurt Konolige[2]
Wolfram Burgard[1]

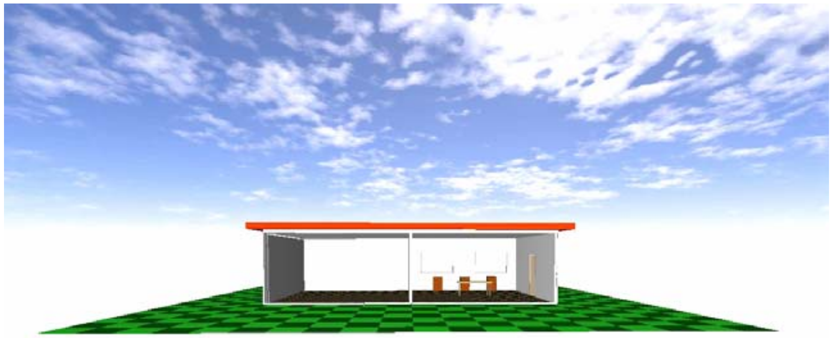
ALBERT-LUDWIGS-UNIVERSITÄT FREIBURG
Willow Garage
STANFORD UNIVERSITY

[1]University of Freiburg
[2]Willow Garage
[3]Stanford University

SA-1

Motivation

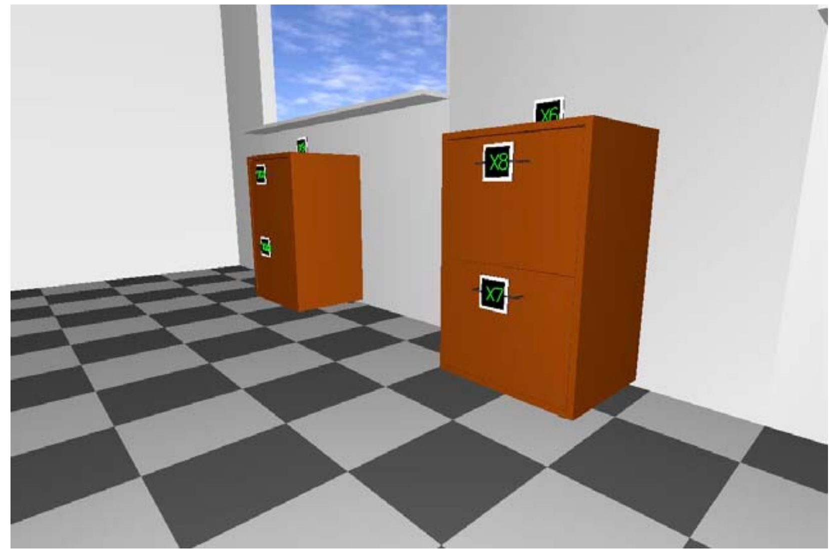
- Domestic environments
- Articulated objects
  - Doors
  - Drawers
- Learning models of objects
- Understanding the spatial movements can improve
  - Perception
  - Manipulation planning
  - Navigation



Problem Formulation

Given: noisy 6D pose observations of *n* rigid object parts

$$x_1 \in \mathbb{R}^6, \ldots, x_n \in \mathbb{R}^6$$



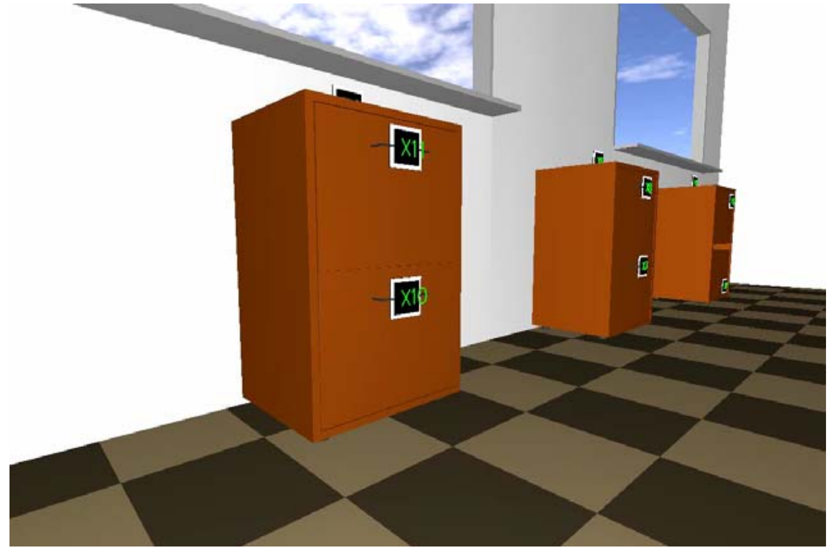
The goal of our approach is to

1. learn models $\mathcal{M}_{ij}$ describing the relationship between two object parts
2. infer the kinematic topology $\mathbb{M}$ of the scene (which object parts are connected in which way)

Training a Model

- Noisy observations
- Consider a sequence of *T* observed relative transforms of $x_i \ominus x_j$, denoted by

$$\mathcal{D}_{ij} = (z_{ij}^1, \ldots, z_{ij}^T)$$



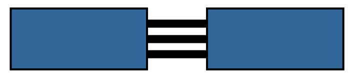
- Find a model $\mathcal{M}_{ij}$ that describes the connection between two object parts (*i* and *j*)
- Maximize the likelihood of the observations given the model

$$\hat{\mathcal{M}}_{ij} = \operatorname{argmax}_{\mathcal{M}_{ij}} p(\mathcal{D}_{ij} \mid \mathcal{M}_{ij})$$

Template Models

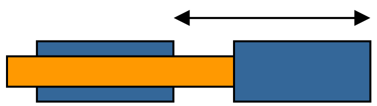
- Rigid model

$\mathcal{M}^{\text{rigid}}$



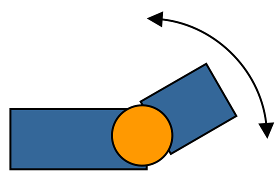
- Prismatic model

$\mathcal{M}^{\text{prismatic}}$



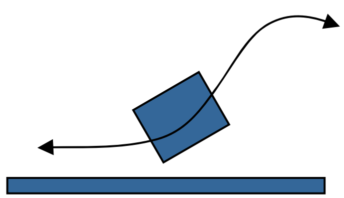
- Rotational model

$\mathcal{M}^{\text{rotational}}$



- Non-parametric model
  (“LLE/GP model”)

$\mathcal{M}^{\text{LLE/GP}}$



Observation Likelihood (1)

- What is the likelihood of an observation given a model?

$$p(z_{ij} \mid \mathcal{M}_{ij}) \; = \; ?$$

- Models for articulated objects typically have a latent action variable *a* (e.g., door opening angle)

$$p(z_{ij} \mid a, \mathcal{M}_{ij}) \propto \exp\left(-||f_{\mathcal{M}_{ij}}(a) \ominus z_{ij}||^2/l\right)$$

- For computing the observation likelihood, we integrate over *a*

$$p(z_{ij} \mid \mathcal{M}_{ij}) = \int_a p(z_{ij} \mid a, \mathcal{M}_{ij}) \; p(a \mid \mathcal{M}_{ij}) \; da$$

Observation Likelihood (2)

- Assumption: all latent actions have the same likelihood

$$p(z_{ij} \mid \mathcal{M}_{ij}) \propto \int_a p(z_{ij} \mid a, \mathcal{M}_{ij}) \, da$$

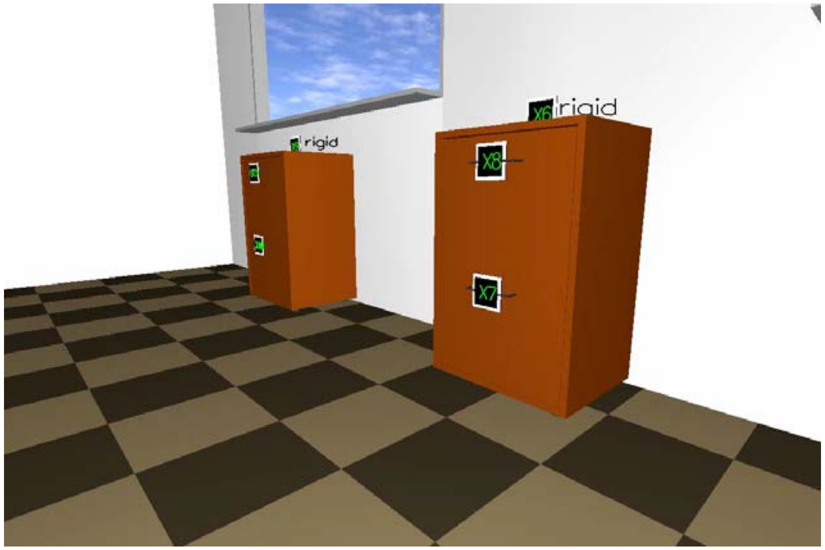
- Integration might be time consuming
- Efficiently estimate value of latent action variable!

$$p(z_{ij} \mid \mathcal{M}_{ij}) \approx \max_a p(z_{ij} \mid a, \mathcal{M}_{ij})$$

- In our settings, this works well in practice for all observed models

Drawer: Moves on Line Segment

- Estimate
  - axis *e* of movement (using PCA)



- Estimate latent action

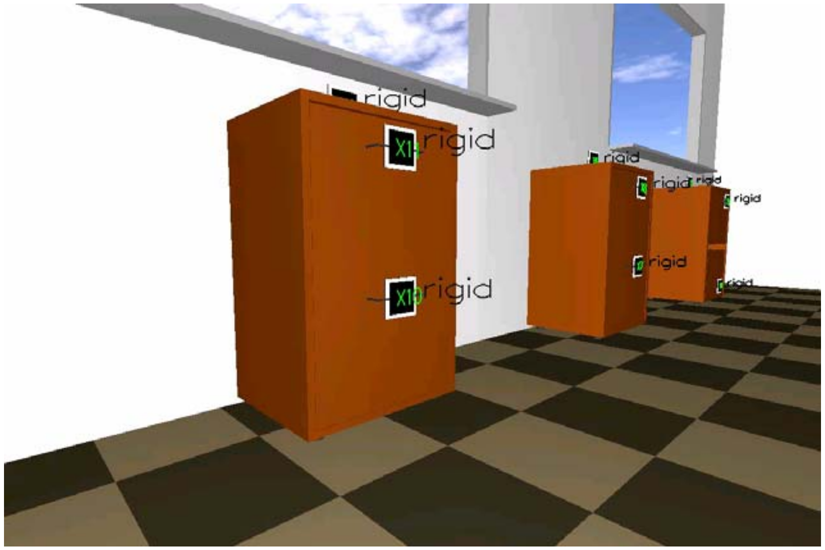
$$\hat{a}_{ij}^{t} = e \cdot trans(z_{ij}^{t} \ominus z_{ij}^{1})$$

- Transformation function:

$$f_{\mathcal{M}_{ij}^{\text{prismatic}}}(a) = z_{ij}^{1} \oplus ae$$

Parametric Model for a Door

- Estimate:
  - axis of rotation *n*
  - center of rotation *c*
  - rigid transform *r*



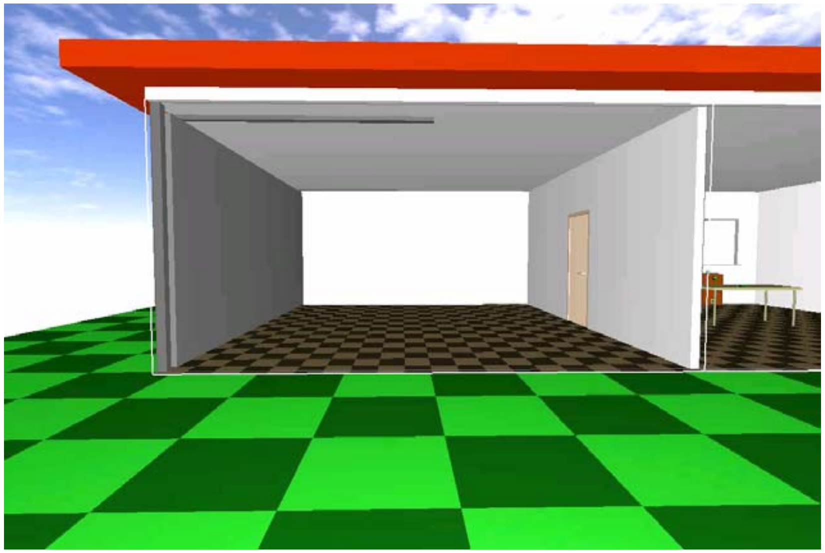
- Estimate latent action

- Transformation function:

$$f_{\mathcal{M}_{ij}^{\text{rotational}}}(a) = [c; n]^T \oplus rot_Z(a) \oplus r$$

Garage door: A Two-link Joint

- Garage door runs in a vertical and a horizontal slider
- Neither rotational, nor prismatic motion
- There are objects which cannot be explained well by “standard” models



LLE / GP: Non-parametric Model

- For a articulation model, we need
  - Estimate the latent action *a*
  - Predict the expected transform

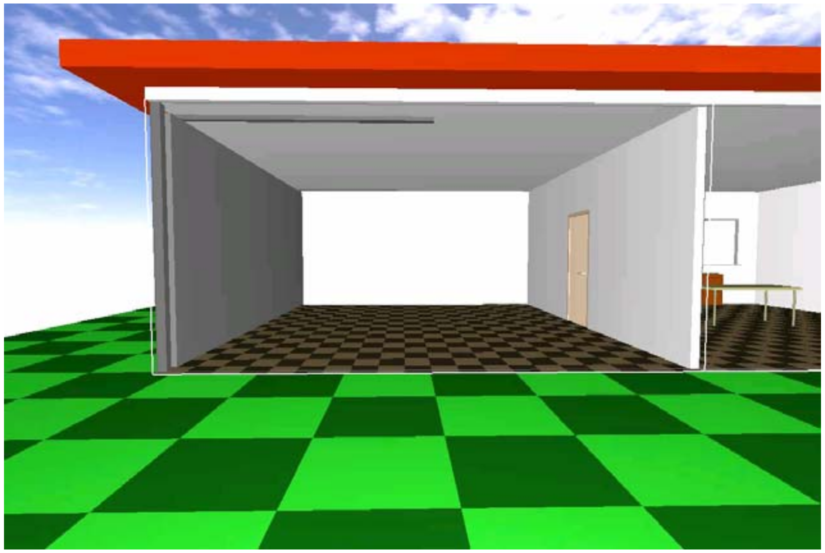
- Assume that the data lies on (or close to) a low dimensional manifold in $\mathbb{R}^6$

- Find latent low dimensional coordinates on the manifold → dimensionality reduction using locally linear embedding (LLE)
- Then learn a Gaussian process regression for the transformation function

$$f(a) = z_{ij}$$

The LLE/GP Model Summary

- Given: 6D obs.
  $z_{ij} \in \mathbb{R}^6$
- Find low-dim. manifold
  $z_{ij} \stackrel{\text{LLE}}{\rightarrow} a_{ij} \in \mathbb{R}^d$
- Non-linear regression
  $(a_{ij}, z_{ij}) \rightarrow \text{GP}$



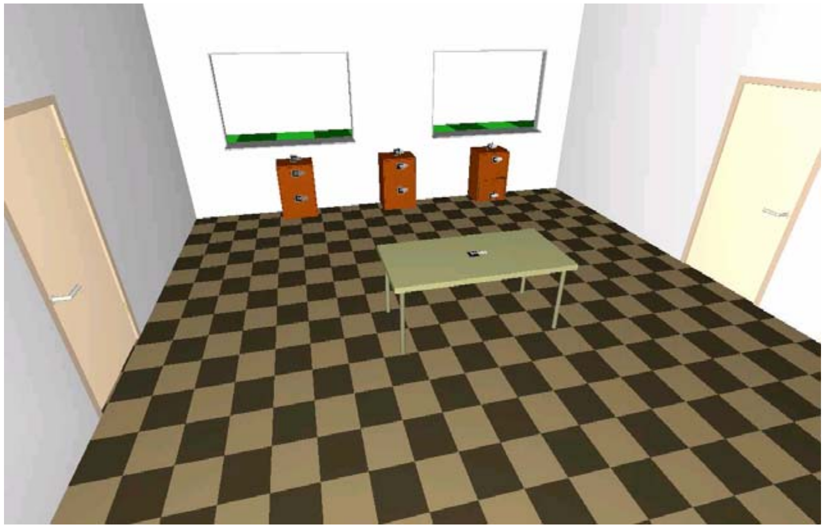
- Result: a mapping from the manifold coordinate system into the observations space

$$a' \stackrel{\text{GP}}{\rightarrow} z' \qquad\qquad f(a) = z_{ij}$$

Higher Dimensional Manifolds

- LLE can find the embedding for any given dimensionality (*d>1*)

- GPs naturally take multi-dimensional input data



⇒ Approach directly applicable

Finding the Topology of the Scene

- Consider all possible models at the same time
- Then select the models that maximize the data likelihood while minimizing the overall complexity

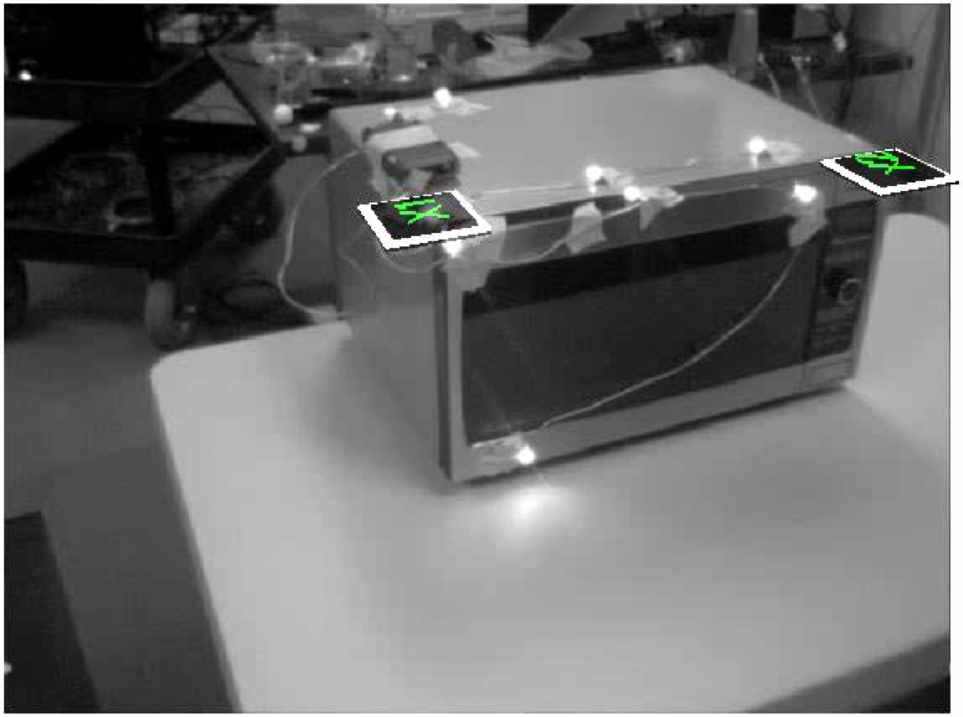
$$\mathrm{cost}_{\mathcal{M}_{ij}^{\mathrm{type}}} = -\frac{1}{\|\mathcal{D}_{ij}^{\mathrm{test}}\|} \log p(\mathcal{D}_{ij}^{\mathrm{test}} \mid \mathcal{M}_{ij}^{\mathrm{type}}) + C(\mathcal{M}_{ij}^{\mathrm{type}})$$

- The model selection and inference of the topology is done by computing a minimum spanning tree

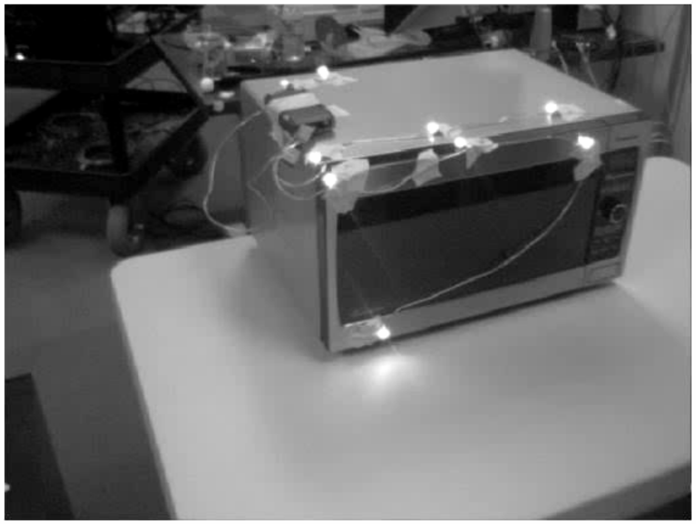
Experiments

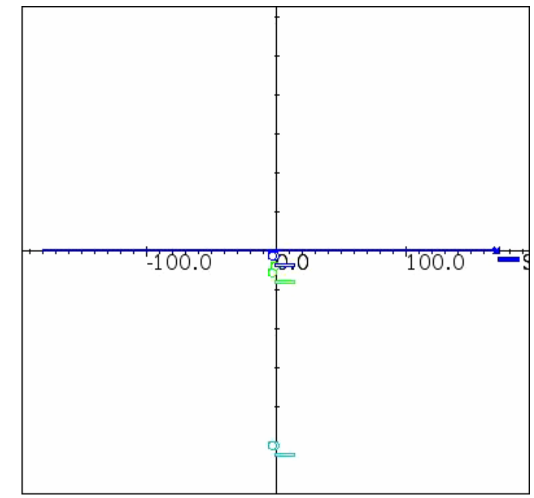
- Microwave door [PhaseSpace]
- Cabinet with two drawers [PhaseSpace]
- Garage door [Simulated]
- Table [ARToolkit]

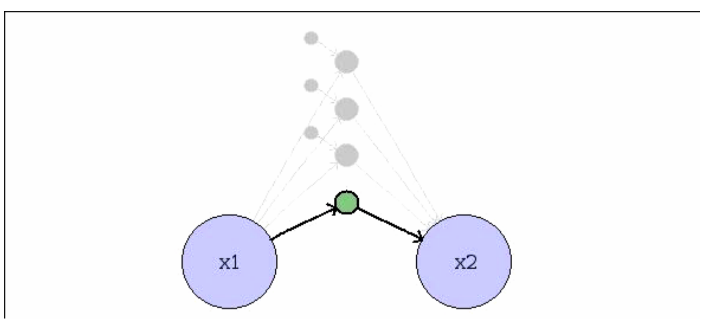
Microwave Door: Observations

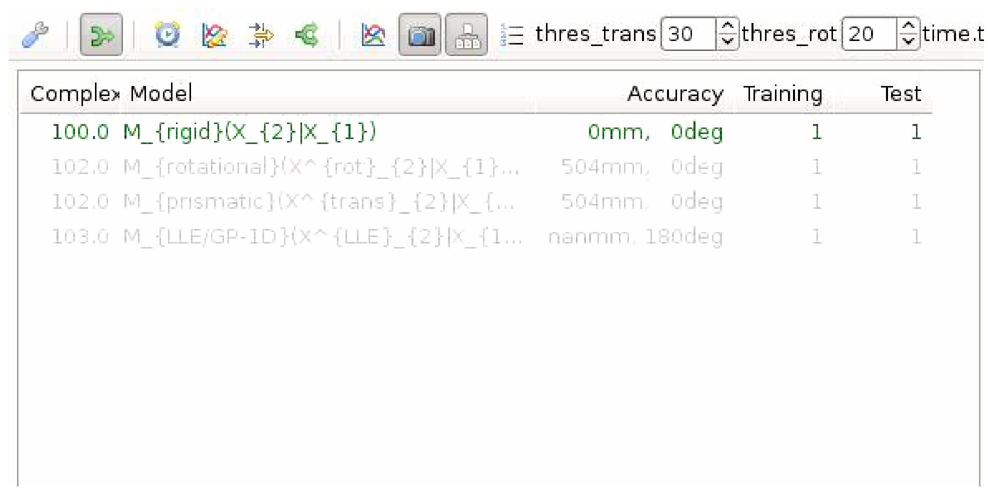


Microwave Door: Learned Model

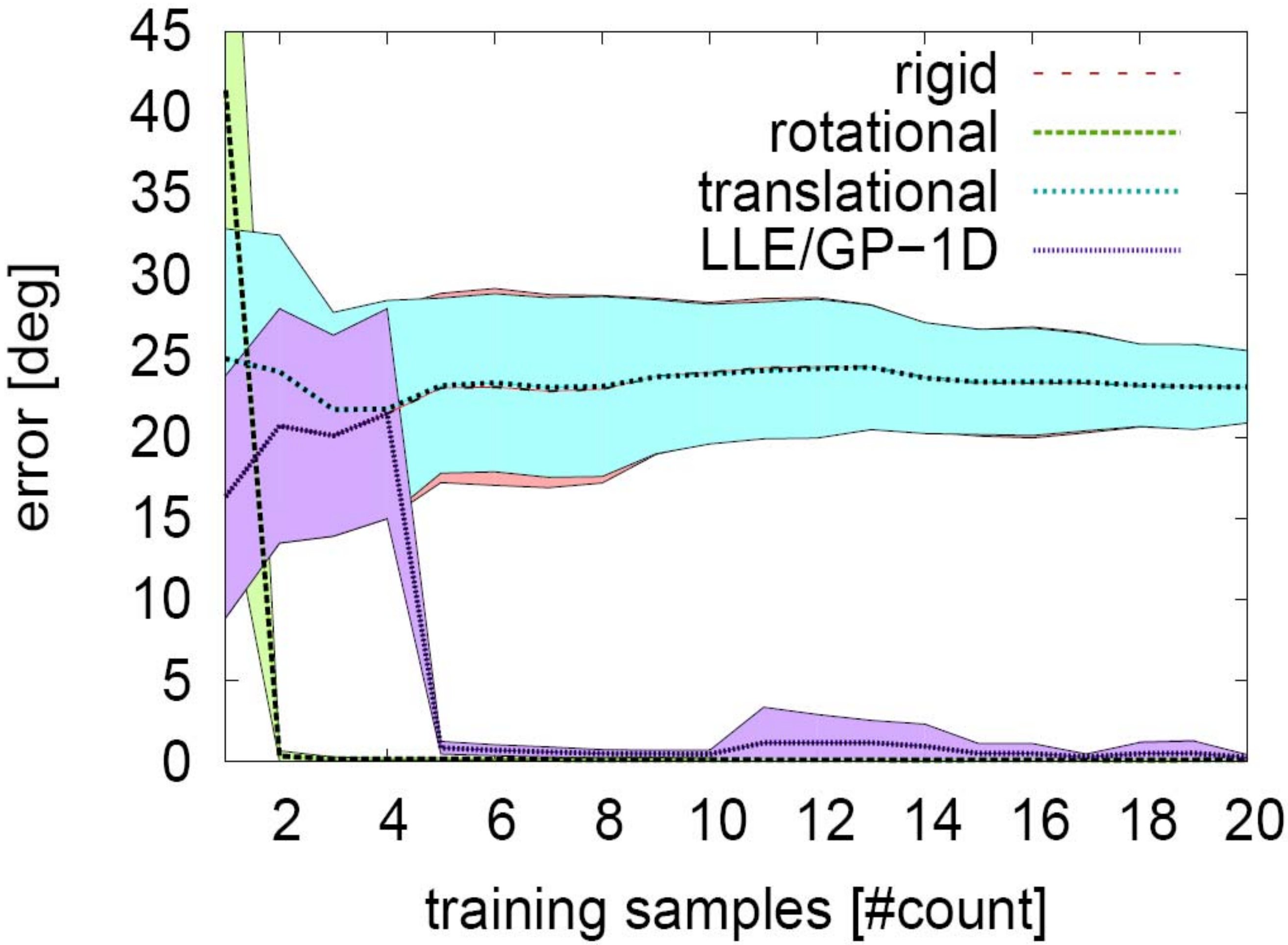




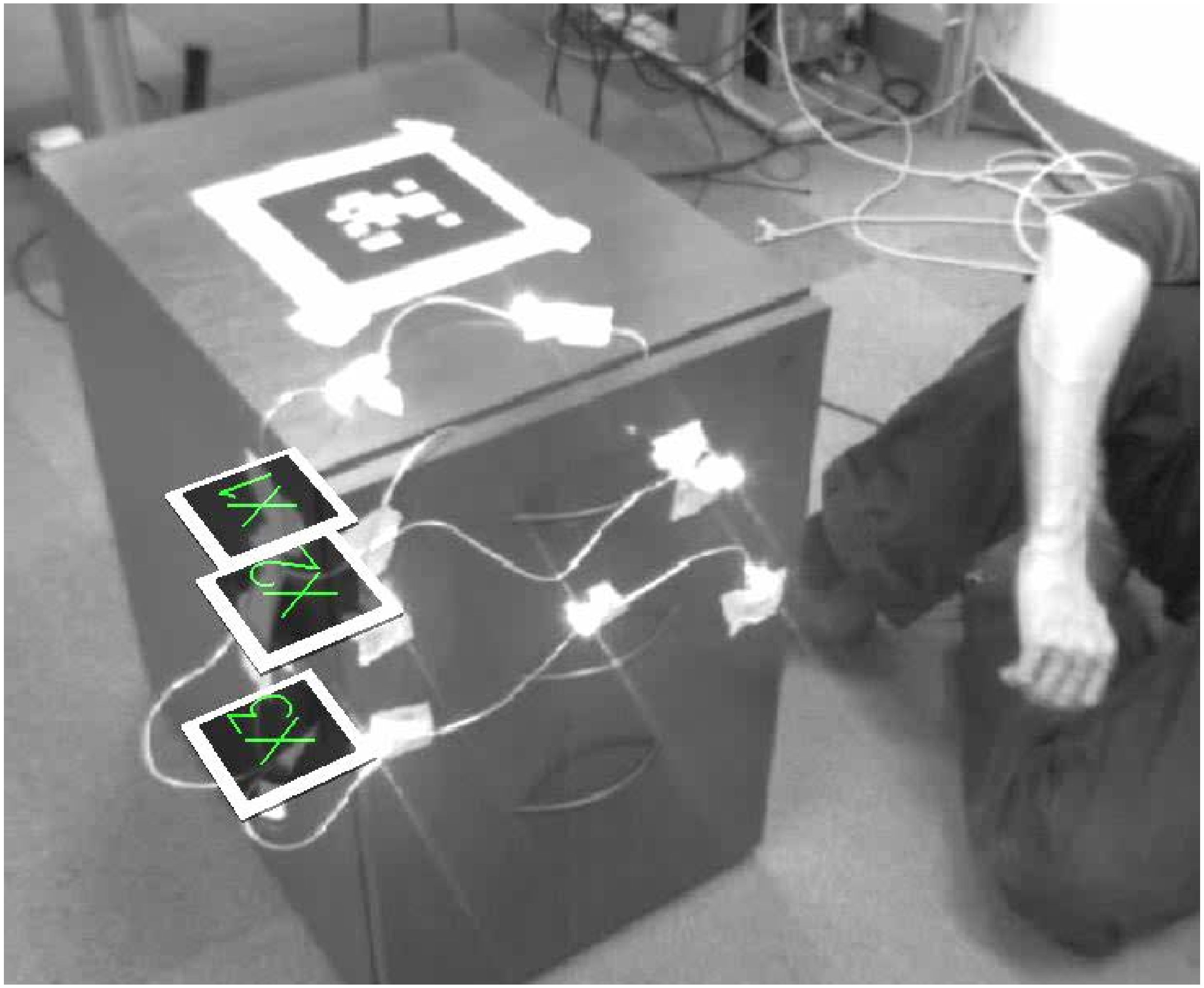




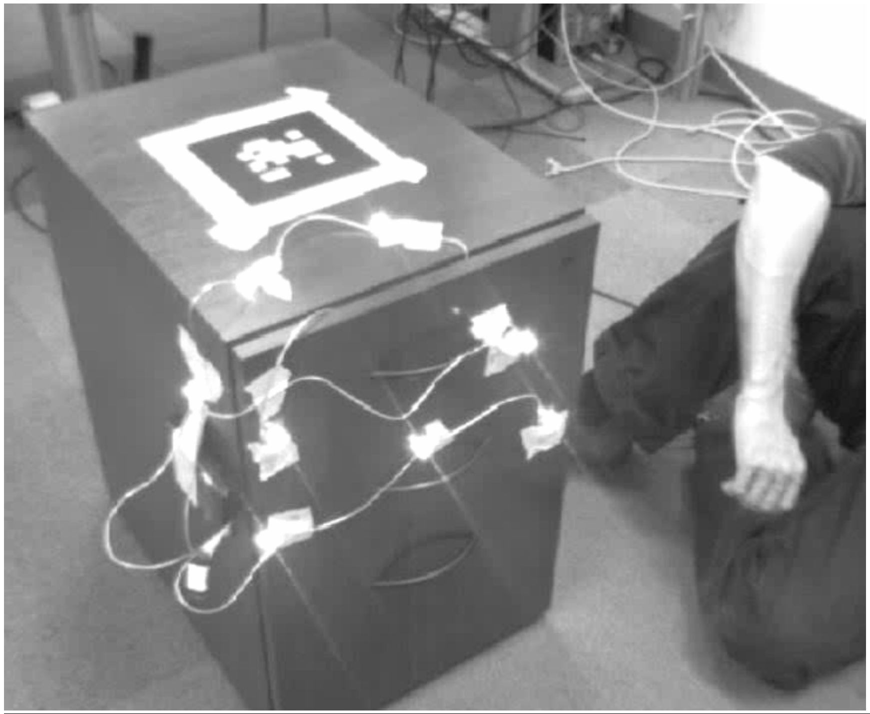
Microwave Door: Error Analysis

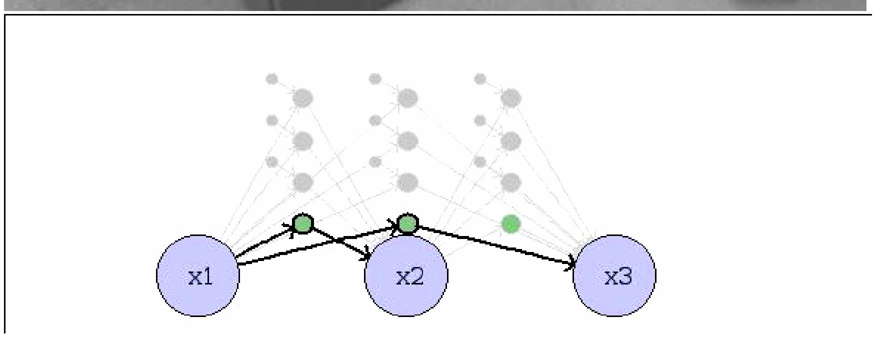


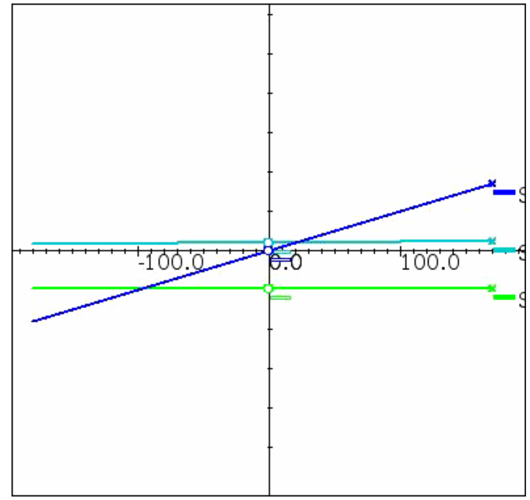
Cabinet with Two Drawers

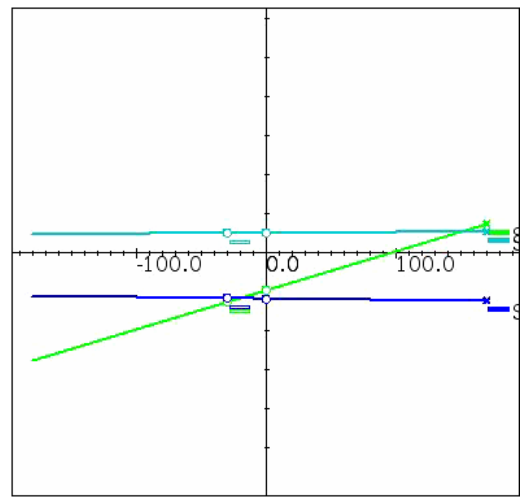


Cabinet: Learned Models

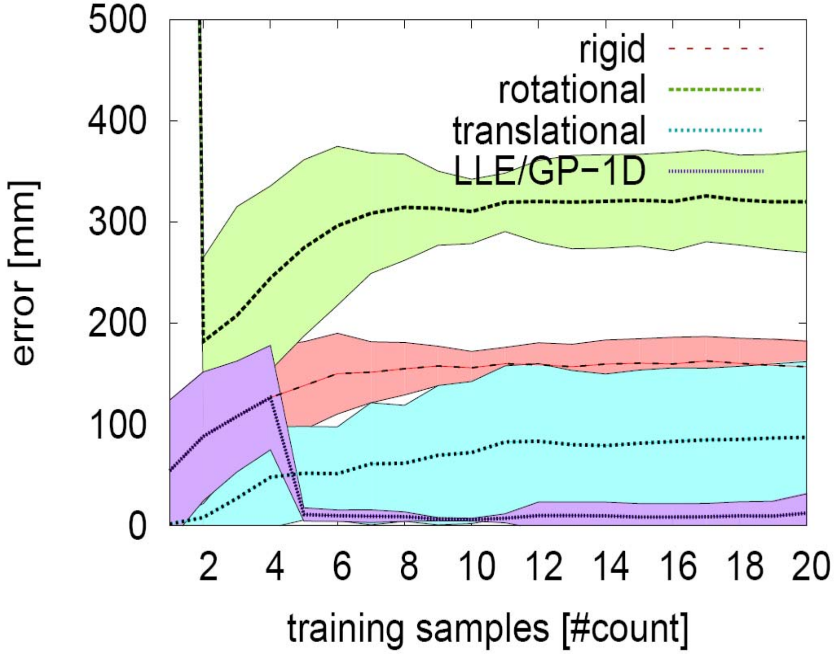


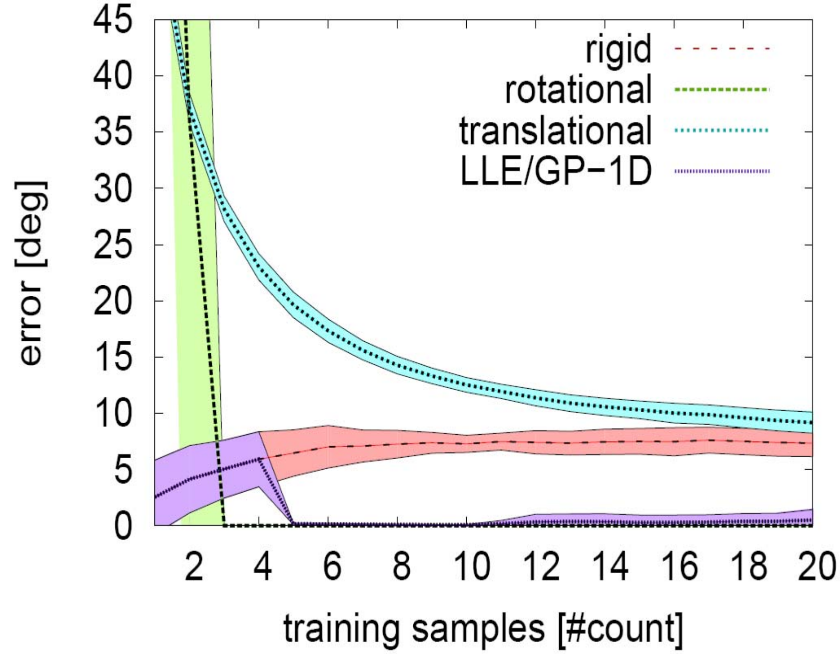






Garage Door: Error Analysis

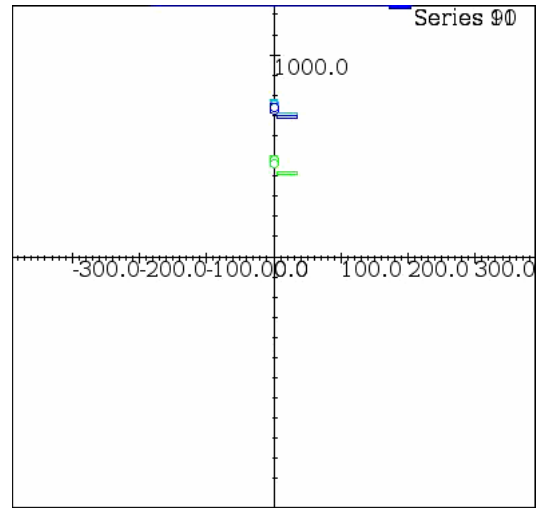


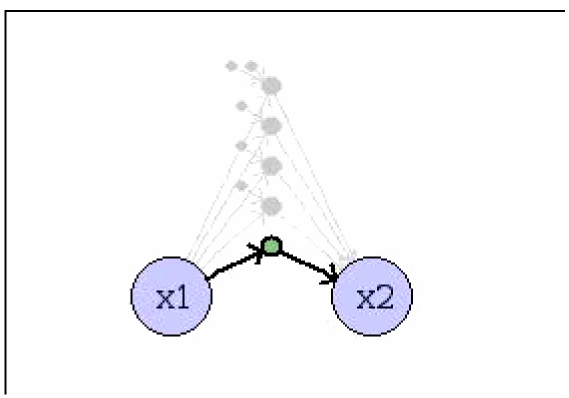


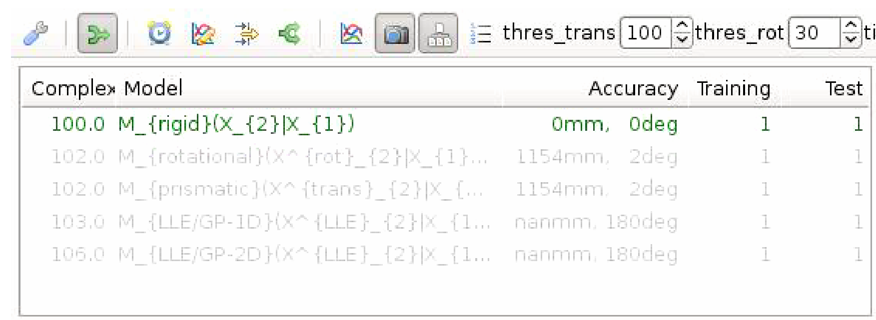
Table

Table

Table


1000.0
-300.0-200.0-100.00.0
100.0 200.0 300.0


x1
x2


thres_trans 100 thres_rot 30
Complex Model Accuracy Training Test
100.0 M_{rigid}(X_{2}|X_{1}) 0mm, 0deg 1 1
102.0 M_{rotational}(X^{rot}_{2}|X_{1}... 1154mm, 2deg 1 1
102.0 M_{prismatic}(X^{trans}_{2}|X_{... 1154mm, 2deg 1 1
103.0 M_{LLE/GP-1D}(X^{LLE}_{2}|X_{1... nanmm, 180deg 1 1
106.0 M_{LLE/GP-2D}(X^{LLE}_{2}|X_{1... nanmm, 180deg 1 1

Conclusions

- Novel approach for learning kinematic models for articulated objects
- Uses a candidate set of parametric and non-parametric models
  - Parametric (rigid, prismatic, rotation)
  - Non-parametric models (LLE/GP)
- Infers the connectivity of the object parts by means of a minimum spanning tree

→ Towards understanding space in domestic environments

Future Work

- Closed kinematic chains
- Pose registration with natural features
- Plan trajectories for door/drawer opening

Thanks for Your Attention